MODELS OF RISK PREFERENCES

RESEARCH IN EXPERIMENTAL ECONOMICS

Series Editors: R. Mark Isaac and Douglas A. Norton

Recent Volumes:

Volume 7: Emissions Permit Experiments, 1999
Volume 8: Research in Experimental Economics, 2001
Volume 9: Experiments Investigating Market Power, 2002
Volume 10: Field Experiments in Economics, 2005
Volume 11: Experiments Investigating Fundraising and Charitable Contributors, 2006
Volume 12: Risk Aversion in Experiments, 2008
Volume 13: Charity With Choice, 2010
Volume 14: Experiments on Energy, The Environment, and Sustainability, 2011
Volume 15: New Advances in Experimental Research on Corruption, 2012
Volume 16: Experiments in Financial Economics, 2013
Volume 17: Experiments in Macroeconomics, 2014
Volume 18: Replication in Experimental Economics, 2015
Volume 19: Experiments in Organizational Economics, 2016
Volume 20: Experimental Economics and Culture, 2018
Volume 21: Experimental Law and Economics, 2022

RESEARCH IN EXPERIMENTAL ECONOMICS VOLUME 22

MODELS OF RISK PREFERENCES: DESCRIPTIVE AND NORMATIVE CHALLENGES

EDITED BY

GLENN W. HARRISON
Georgia State University, USA

and

DON ROSS
University College Cork, Ireland
University of Cape Town, South Africa
Georgia State University, USA

SERIES EDITORS

R. MARK ISAAC
Florida State University, USA

and

DOUGLAS A. NORTON
Florida State University, USA

United Kingdom – North America – Japan
India – Malaysia – China

Emerald Publishing Limited
Emerald Publishing, Floor 5, Northspring, 21-23 Wellington Street, Leeds LS1 4DL.

First edition 2023

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-83797-269-2 (Print)
ISBN: 978-1-83797-268-5 (Online)
ISBN: 978-1-83797-270-8 (Epub)

ISSN: 0193-2306 (Series)

CONTENTS

This page intentionally left blank

ABOUT THE EDITORS

Glenn W. Harrison is University Distinguished Professor, C.V. Starr Chair of Risk Management & Insurance and Director of the Center for the Economic Analysis of Risk, Maurice R. Greenberg School of Risk Science, J. Mack Robinson College of Business, Georgia State University. His current research spans risk management & perception, experimental economics, behavioral econometrics, behavioral welfare economics, and development economics.

Don Ross is Professor in the School of Society, Politics, and Ethics at University College Cork, Professor in the School of Economics at the University of Cape Town, and Program Director for Methodology at the Centre for the Economic Analysis of Risk at Georgia State University. His current research focuses on experimental studies and theoretical modeling of risk and time preferences in humans and other animals, gambling disorders and policy, economic methodology, the economics of road transport networks in Africa, and the metaphysical implications of science.

This page intentionally left blank

LIST OF CONTRIBUTORS

Glenn W. Harrison	Maurice R. Greenberg School of Risk Sciences and Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, USA; School of Economics, University of Cape Town, South Africa
Morten I. Lau	Department of Economics, Copenhagen Business School, Copenhagen, Denmark; Department of Economics, Durham University, Durham, UK
Brian Albert Monroe	School of Economics, University of Cape Town, South Africa
Don Ross	School of Society, Politics and Ethics, University College Cork, Ireland; Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, USA; School of Economics, University of Cape Town, South Africa
J. Todd Swarthout	Department of Economics, Andrew Young School of Policy Studies, Georgia State University, USA; Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, USA
Nathaniel T. Wilcox	Department of Economics, Walker College of Business, Appalachian State University, Boone, USA
Hong Il Yoo	School of Business and Economics, Loughborough University, Loughborough, UK
Hongming Zhao	Department of Economics, Durham University, Durham, UK

This page intentionally left blank

INTRODUCTION

Glenn W. Harrison and Don Ross

The insights of behavioral economics have directly influenced policy in recent decades, and have started to influence the way in which welfare economics is used to design and evaluate policy. The need to provide guidance for policy design premised on behavioral patterns, and evaluate existing or new policy, has generated a "derived demand" for better descriptive models of behavior as well as better ways to understand relevant normative standpoints. This volume brings together contributions to that challenge, with a focus on behavior with respect to risky choices.

In Chapter 1, Glenn W. Harrison and Don Ross develop a philosophically consistent approach to behavioral welfare economics. They call their approach the Quantitative Intentional Stance (QIS), based on the Intentional Stance (IS) of Dennett (1971, 1987). For Dennett, there are complementary styles of scientific explanation that he calls "stances," and the one he calls the IS is the stance most generally applicable in cognitive and behavioral science. It also happens to be close to how economists have historically constructed their explanations and models, which makes it an ideal basis for linking the modeling of descriptive behavior toward risk with the modeling of normative evaluations of this behavior. The observer, or normative policy-maker, forms priors about the (risk) preferences and (risk) perceptions of the agent being evaluated, and the choice environment the agent faces, using the best models that her evidence allows. Then the observer forms priors about how those preferences and beliefs should affect choices over risky prospects, and contrasts those expected choices with observed choices. Sometimes the agent makes the choices "expected" by the observer given her priors over preferences, beliefs, and how they should affect choices, and sometimes they are different. The former then constitute welfare gains for the agent, and the latter constitute welfare losses for the agent.[1]

All of this imposes a burden on the economist to generate rigorous models of risk preferences and risk perceptions, which is why there is the derived demand for more care being given to the descriptive side of the structural modeling of risk.

Models of Risk Preferences: Descriptive and Normative Challenges
Research in Experimental Economics, Volume 22, 1–6

ISSN: 0193-2306/doi:10.1108/S0193-230620230000022007

Harrison and Ross contrast their QIS with existing approaches toward behavioral welfare economics in the economics literature, some of which err because they conflate the economist's technical concept of welfare with wider understandings of well-being that do *not* admit of technical specification. This raises the question of how policy-makers, and agents themselves, should regard the normative force of the economist's technical analysis. Answers to this question are highly context-sensitive, but some guidelines are provided for identifying frequently relevant contextual features. Harrison and Ross also directly answer a wide range of questions about the QIS that arise from those looking for easier solutions to the challenges of behavioral welfare economics, explaining the dangers of various short-cuts.

A first step at reconsidering the structural descriptive evidence about risk preferences is taken in Chapter 2 by Nathaniel T. Wilcox. He considers a popular meme, according to which everyone exhibits a specific type of bias in the way they process risky probabilities: that they overweight the extreme outcomes, and underweight outcomes in between. This presumption arose from early statements of Original Prospect Theory by Kahneman and Tversky (1979), and in later statements of Cumulative Prospect Theory (CPT) by Tversky and Kahneman (1992). To an economist, a model of risk preferences does not normally come with specific functional forms or parameter values for those functions as parts of the model, but in some parts of psychology that is common. Many people believe that CPT *requires* this particular mix of optimism and pessimism in the face of risk. Economists take the view that the general validity of the way that CPT explains risky choices should not stand or fall on particular patterns of probability weighting.

In that spirit, Wilcox elegantly lets the data speak on the matter, by careful experimental design as well as careful econometric evaluation of the observed data. He designs a battery of choice tasks over risky prospects that allow a wide range of utility functions and probability weighting functions to play a role in explaining choices for each individual, and ensures that the resulting data should have high power for the inferences to then come. And in the econometric evaluation he completely avoids taking a stance on the functional form of either the utility function or the probability weighting function. He also avoids pooling behavior over individuals. Even with conditioning on observable characteristics of individual subjects, pooling prevents the data from revealing the full heterogeneity of behavior at the individual level.

The methodological drive to model behavior at the individual level is more significant than it might seem. What if one third of the subjects in a sample overweighted the worst outcomes in the battery, one third overweighted none of the outcomes, and one third overweighted the best outcomes in the battery? Pooling data from these three cohorts, and assuming their choices are all determined by one specific probability weighting function, the observer would be led to find support for the meme when, by construction, there is none at the individual level. Of course, one might just wag a finger at such inferences, chiding that the average of a distribution of choices by many individuals should never be assumed to apply

exactly to *every* choice by *every* individual, or even *any* choice by *any* individual … except that one would be wagging fingers at a vast literature! The heterogeneity that Wilcox documents with respect to probability weighting is convincing testimony to the dangers of pooled estimation.

In Chapter 3, Glenn W. Harrison and J. Todd Swarthout pick up these themes, and set about evaluating what fraction of risky choices are best characterized descriptively by CPT in a controlled lab setting, where the initial empirical insights underlying CPT arose. In contrast to Wilcox in Chapter 2, they assume popular, flexible functional forms from the literature. Consistent with Wilcox, they stress the need to estimate risk preferences at the level of the individual. Hence, they design and conduct experiments with hundreds of subjects, each making a large number of choices over risky prospects that allow CPT to play a role. Their approach is to estimate several models of risk preferences for each individual: an Expected Utility Theory (EUT) model, a Dual Theory (DT) model, a Rank-Dependent Utility (RDU) model, a Disappointment Aversion (DA) model, and a CPT model. They allow the DT and RDU models to employ a particularly flexible probability weighting function, mindful in their priors of the heterogeneity in probability weighting that Wilcox demonstrates so thoroughly. And their choice battery includes many choices in a gain frame, a loss frame, and a mixed gain-loss frame, as required to identify the core parameters of CPT. Of course, although not "of course" to much of the extant CPT literature, all choices involved real monetary rewards.

Their striking finding is that CPT performs poorly when the risky prospects are set up to give it the most favorable chance of explaining behavior. The clear modal model accounting for behavior is RDU, with EUT solidly in second place. The root of the poor performance of CPT is easy to identify: local asset integration by subjects when faced with losses out of an endowment, whether or not that endowment was provided by the experimenter or earned in some prior task. If an individual has a $60 endowment and faces a prospect of losing $20 from that, does she evaluate the net gain of $40 or does she evaluate the gross loss of $20? CPT assumes the latter, at least under plausible assumptions about what "the" reference point for this subject is. But if she locally integrates the loss of $20 with the endowment of $60 and just evaluates the net gain of $40, CPT is "dead in the water" in terms of what differentiates it from RDU and EUT. Here, it is important to recognize that RDU and EUT are *not* nested within CPT, as commonly and casually claimed.

The data of Harrison and Swarthout, as well as their detailed cautions about short-cut models that allegedly find evidence for the beloved loss aversion parameter of CPT, should provide a rich starting point for consideration of further evaluations of CPT. Apart from DA, existing models of "endogenous reference points" provide no useful *operational* guidance for estimation. So perhaps we need more operational guidance from theory on what the reference point of CPT might be. And even with the large battery of choices each subject made, estimating all of the moving parts of CPT poses serious econometric challenges which Harrison and Swarthout urge others to take up with Bayesian Hierarchical

Models that use pooled behavior solely to provide weakly informative priors with which to tease apart those moving parts at the level of the individual. In our own work, we simply do not bother with CPT any more, since it is a needless and ill-motivated distraction for the behavioral welfare economist. If advocates of CPT ever take the incentivized data and econometric demands more seriously, we are of course open to reconsider that judgment.

In Chapter 4, Morten I. Lau, Hong Il Yoo, and Hongming Zhao consider a different aspect of risk preferences if they are to be used to make normative evaluations, their stability over time. They start from a recognition that "stability" can mean several things. It might mean unconditional stability over time, even as characteristics and "states of nature" change, or it might mean conditional stability given values for those characteristics and states.[2] And it might mean that some population exhibits stability over time, even if individuals do not, or it might mean that individuals exhibit stability over time. They use panel data from influential longitudinal lab experiments previously applied to estimate CPT risk preferences by Glöckner and Pachur (2012) and Murphy and ten Brincke (2018). Their econometric approach extends a random coefficient approach to estimation of risk preferences developed by Harrison, Lau, and Yoo (2020) to consider CPT, rather than just EUT and RDU.

CPT posits three possible "pathways" to explain the risk premium. One is aversion to variability of outcomes, one is pessimism with respect to better outcomes, and one is a special aversion to losses. One of the core findings of Lau, Yoo, and Zhao is that none of the three pathways to the risk premium in CPT seems to be stable over time, no matter what definition of temporal stability is used. A second core finding is that temporal stability varies with the model of risk preferences adopted. For example, for their data and models, they conclude that EUT exhibits temporal stability even if CPT does not, even when applied to the same data. One can always point to cautions about sample sizes, payoff scales, econometric specifications, and other methodological attributes of the comparisons, and Lau, Yoo, and Zhao do this carefully. But it is apparent that temporal stability of some kind[3] is needed in order to undertake normative evaluations of policies to be applied in the future, so further attention to these issues is needed.

Finally, in Chapter 5, Brian Albert Monroe takes up a subtlety in the application of the QIS of surprising significance for welfare evaluations. In an early application of the QIS by Harrison and Ng (2016) subjects were asked to make a series of binary choices over risky lotteries, to allow inferences about the risk preferences of each individual. Those inferences were, in turn, to be used as priors in the normative evaluation of insurance purchase decisions that the same subjects made in a separate task. A key step in the evaluation was to initially determine if the risk preferences of each individual were better characterized by EUT or by RDU, and then to use the estimates[4] for EUT *or* RDU for that individual when evaluating the insurance choices by that individual. An individual was assumed to be characterized by EUT unless their choices over the risky lotteries exhibited statistically significant evidence of probability weighting, using appropriate tests and various significance levels. In large part, this initial step of "typing" the individual as EUT or RDU was undertaken to make the point that the normative

evaluation of insurance choices depends on the *type* of risk preference as well as the *level* of risk aversion.

Monroe explains that there are two potential problems with this approach. The first problem is that subjects who are judged to be better characterized as EUT decision-makers could just be extremely noisy RDU decision-makers. And there is no reason to expect that the "noise" in question here affects inferences in some simple additive, linear manner that might wash out when making normative evaluations. The second problem is that declaring somebody to be better characterized as an EUT decision-maker, and then using their EUT estimates for normative evaluation, is actually saying that they are a RDU decision-maker with *exactly* estimated parameters for their probability weighting function.[5] This is not the same thing as saying that they are "sufficiently noisy" in terms of probability weighting that one cannot reject a null hypothesis that they exhibit no probability weighting at all.

What is remarkable, and skillfully demonstrated in numerical simulations by Monroe, is that these seemingly subtle steps in the descriptive characterization of risk preferences can have a significant impact on normative evaluation. The impacts are most pronounced on the inferred size of the welfare gain or loss, rather than the sign of the welfare effect, but that is not the main methodological point. The important lesson from Monroe, already adopted in later studies[6] with reference to his arguments, is to just use the RDU model for every subject when undertaking the normative evaluations of welfare. One can still usefully use the classification of risk preference type, by whatever means one wants, to help understand the sources of welfare gains and losses, but one should not use the special case of EUT when an individual is better characterized for *normative* purposes as RDU and EUT is nested in RDU. The clear exception here is if the policy-maker or analyst has a well-motivated and explicit reason to maintain EUT as the normative metric for evaluating choices, in which case every subject's risk preferences should be characterized by the EUT model, even if the RDU model does a better job for *descriptive* purposes.

Taken as a whole, the contributions here constructively make the case that structural insights into risk preferences are needed in order to make rigorous normative evaluations of risky choices. They also clearly flag remaining challenges in undertaking these normative evaluations.

NOTES

1. The key to the confident statement here is that if the QIS has been rigorously applied in a Bayesian inferential procedure, there is nothing more than anyone could relevantly know about the choice situation, at least in isolation from possible general equilibrium effects as unintended consequences.
2. See Andersen, Harrison, Lau, and Rutström (2008).
3. In fact it is temporal *predictability* of risk preferences that is needed, but that can just be viewed as stability conditional on evolving characteristics that affect risk preferences.
4. By "estimates" we mean point estimates as well as estimates of standard errors and covariances, which were incorporated in the normative evaluation using bootstrap procedures.

5. And, of course, that those exact values imply zero probability weighting consistent with EUT.

6. For example, Gao, Harrison, and Tchernis (2023) and Harrison, Morsink, and Schneider (2022).

REFERENCES

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Lost in state space: Are preferences stable? *International Economic Review*, *49*(3), 1091–1112.

Dennett, D. (1971). Intentional systems. *The Journal of Philosophy*, *68*(4), 87–106.

Dennett, D. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

Gao, X. S., Harrison, G. W., & Tchernis, R. (2023). Behavioral welfare economics and risk preferences: A Bayesian approach. *Experimental Economics*, *26*, 273–303.

Glöckner, A., & Pachur, T. (2012). Cognitive models of risky choice: Parameter stability and predictive accuracy of prospect theory. *Cognition*, *123*(1), 21–32.

Harrison, G. W., Lau, M. I., & Yoo, H. I. (2020). Risk attitudes, sample selection and attrition in a longitudinal field experiment. *Review of Economics and Statistics*, *102*, 552–568.

Harrison, G. W., Morsink, K., & Schneider, M. (2022). Literacy and the quality of index insurance decisions. *Geneva Risk and Insurance Review*, *47*, 66–97.

Harrison, G. W., & Ng, J. M. (2016). Evaluating the expected welfare gain of insurance. *Journal of Risk & Insurance*, *83*(1), 91–120.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Murphy, R. O., & ten Brincke, R. H. W. (2018). Hierarchical maximum likelihood parameter estimation for cumulative prospect theory: Improving the reliability of individual risk parameter estimates. *Management Science*, *64*(1), 308–326.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representations of uncertainty. *Journal of Risk & Uncertainty*, *5*: 297–323.

CHAPTER 1

BEHAVIORAL WELFARE ECONOMICS AND THE QUANTITATIVE INTENTIONAL STANCE

Glenn W. Harrison and Don Ross

ABSTRACT

Behavioral economics poses a challenge for the welfare evaluation of choices, particularly those that involve risk. It demands that we recognize that the descriptive account of behavior toward those choices might not be the ones we were all taught, and still teach, and that subjective risk perceptions might not accord with expert assessments of probabilities. In addition to these challenges, we are faced with the need to jettison naive notions of revealed preferences, according to which every choice by a subject expresses her objective function, as behavioral evidence forces us to confront pervasive inconsistencies and noise in a typical individual's choice data. A principled account of errant choice must be built into models used for identification and estimation. These challenges demand close attention to the methodological claims often used to justify policy interventions. They also require, we argue, closer attention by economists to relevant contributions from cognitive science. We propose that a quantitative application of the "intentional stance" of Dennett provides a coherent, attractive and general approach to behavioral welfare economics.

Keywords: Behavioral economics; welfare economics; behavior toward risk; revealed preference; normative economics; cognitive science

Models of Risk Preferences: Descriptive and Normative Challenges
Research in Experimental Economics, Volume 22, 7–67

ISSN: 0193-2306/doi:10.1108/S0193-230620230000022001

One theme lies at the heart of a rigorous evaluation of policy using the insights of behavioral welfare economics: how to judge if some policy is encouraging good choices or bad choices. One approach, which drives the nudge movement and some randomized evaluations, is to assume that judgment away, and simply assert that some change in an *observable* must be good regardless of special features of the preferences and beliefs of individuals. Isn't it obvious that people should save more, eat fewer fatty foods, drink less wine, and take up insurance? The only appropriate answer to this for an economist is "no." Demand for these behaviors depends on preferences and beliefs, and hence the expected consumer surplus (CS) they deliver is also conditional on preferences and beliefs. Further, no realistic economics that recognizes heterogeneity of preferences and beliefs, and identifies a welfare enhancing policy conditional on those preferences and beliefs, tells us that such policies would be welfare-enhancing *on average* if applied *unconditionally*.

Revealed preference theory (RPT), the various axiomatizations of which are reviewed by Chambers and Echenique (2016), minimizes the inferential gap between agents' observed behavior and ascribed goals by treating preferences as summaries of actual choices. As emphasized by Binmore (2009), under RPT it is simply an *error* to regard preferences as *causes* of choices, or as *explaining* them. The motivation for this is not philosophical commitment to behaviorism about "mental states," but reflects the value to the economist, for the sake of theoretical power, of treating all processes that generate common utility outcomes for an agent as an equivalence class. This raises problems for normative analysis, however, because it seems to leave no room for treating any individual choices as reflecting *error*. In effect, every choice that the agent makes is assumed to make her no worse off. Such an idealization makes it impossible *a priori* for the economist to offer any advice about prospective choices that would be potentially superior with respect to the agent's subjective welfare.

Behavioral welfare economics seeks to bridge this gap. One obvious way to do this, which is common in the literature, is to simply abandon RPT, and directly model preferences as hypothesized psychological states that cause choices. The theorist is then free to suppose that different psychological processes have varying welfare consequences, and that an agent can profit from the economist's advice whenever her choices result from less-than-optimal such processes. But this approach has major costs. It makes welfare economics conditional on strong psychological assumptions that must be expected to vary from case to case, even across choices that are identical with respect to outcomes. Thus, it directly undermines the kinds of generalizations economists care about both descriptively and normatively. The psychological assumptions in question are often, necessarily, pure conjectures. Where they are based on psychological experiments, they are hostage to inferential methods in psychology that are heavily criticized, have produced many post-publication retractions by journals, and are rejected by most economists when applied in their own discipline (Ortmann, 2021; Yarkoni, 2022). Appeals to hypothesized brain states based on neuroimaging studies involve yet more egregious violations of econometric best practice with respect to both identification and estimation (Harrison, 2008; Harrison & Ross, 2020). Worst of all